



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Evolving building blocks of rhythm

Citation for published version:

Ravignani, A, Thompson, B, Grossi, T, Delgado, T & Kirby, S 2018, 'Evolving building blocks of rhythm: How human cognition creates music via cultural transmission', *Annals of the New York Academy of Sciences*, vol. 1423, no. 1, pp. 176-187. <https://doi.org/10.1111/nyas.13610>

Digital Object Identifier (DOI):

[10.1111/nyas.13610](https://doi.org/10.1111/nyas.13610)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Annals of the New York Academy of Sciences

Publisher Rights Statement:

This is the peer reviewed version of the following article: Ravignani, A., Thompson, B., Grossi, T., Delgado, T. and Kirby, S. (2018), Evolving building blocks of rhythm: how human cognition creates music via cultural transmission. *Ann. N.Y. Acad. Sci.* doi:10.1111/nyas.13610, which has been published in final form at <http://onlinelibrary.wiley.com/doi/10.1111/nyas.13610/full>. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



ANNALS *of* THE NEW YORK ACADEMY OF SCIENCES

Evolving building blocks of rhythm: How human cognition creates music via cultural transmission

Journal:	<i>Ann NY Acad Sci</i>
Manuscript ID	annals-1748-002.R1
Manuscript Type:	Concise Original Report
Date Submitted by the Author:	n/a
Complete List of Authors:	Ravignani, Andrea; Max-Planck-Institut fur Psycholinguistik, Language and Cognition Department; Sealcentre Pieterburen, Veterinary & Research Department Thompson, Bill; Max-Planck-Institut fur Psycholinguistik, Language and Cognition Department Grossi, Thomas; University of Edinburgh, Centre for Language Evolution, School of Philosophy, Psychology and Language Sciences Delgado, Tania; University of California San Diego , Department of Cognitive Science Kirby, Simon; University of Edinburgh, Centre for Language Evolution, School of Philosophy, Psychology and Language Sciences
Keywords:	evolution of music, rhythm, timing, cultural transmission, biomusicology

SCHOLARONE™
Manuscripts

Full Title (no longer than 15 words): **Evolving building blocks of rhythm: How human cognition creates music via cultural transmission**

Manuscript submitted to *Annals of NYAS* - Issue Neuromusic 2017, invited submission

Authors' names:

Andrea Ravignani^{a,b,c,*}, Bill Thompson^{a,b,*}, Thomas Grossi^d, Tania Delgado^{d,e}, Simon Kirby^d

Author's affiliations:

^a Artificial Intelligence Lab, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium.

^b Language and Cognition Department, Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD Nijmegen, The Netherlands.

^c Research Department, Sealcentre Pieterburen, Hoofdstraat 94a, 9968 AG Pieterburen, The Netherlands.

^d Centre for Language Evolution, School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Edinburgh EH8 9AD, UK.

^e Department of Cognitive Science, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0515, USA.

* Equal contributions.

Corresponding author's contact information:

Andrea Ravignani, Research Department, Sealcentre Pieterburen, Hoofdstraat 94a, 9968 AG Pieterburen, The Netherlands

Phone: 0031 650474647

E-mail: andrea.ravignani@gmail.com

Short Title (<50 characters): **Rhythm cognition via cultural evolution**

Keywords (4-6 keywords separated by semicolons):

Evolution of music; rhythm; timing; cultural transmission; biomusicology; Bayesian model

Abstract (<200 words): Why does musical rhythm have the structure it does? Musical rhythm, in all its cross-cultural diversity, exhibits commonalities across world cultures. Traditionally, music research has been split in two fields. Some scientists focused on musicality, namely the human biocognitive predispositions for music, with an emphasis on cross-cultural similarities. Other scholars investigated music, seen as cultural product, focusing on the large variation in world musical cultures. Recent experiments found deep connections between music and musicality, reconciling these opposing views. Here we address the question of how individual cognitive biases affect the process of cultural evolution of music. Data from two experiments is analyzed using two complementary techniques. In the experiments, participants hear drumming patterns and imitate them. These patterns are then given to the same or another participant to imitate. The structure of these - initially random - patterns is tracked along experimental 'generations'. Frequentist statistics show how participants' biases are amplified by cultural transmission, making drumming patterns more structured. Structure is achieved faster in transmission within, rather than between, participants. A Bayesian model approximates the motif structures participants learned and created. Our data and models suggest that individual biases for musicality may shape the cultural transmission of musical rhythm.

Main text (no more than 3500 words, 2–4 figures and/or tables): **Word count:** 3427

Introduction

The ability to produce music is a human behaviour found universally in all human cultures. Though within the scientific study of music, there is a long-standing division between those who are concerned with music as a cultural product, versus those who focus primarily on music and the mind¹. Traditionally, music researchers distinguish *music* from *musicality*²: music is a cultural product, while musicality is the human biological machinery used to produce and process music. At one extreme, cultural anthropologists and field researchers have historically focused on *music* as a cultural product^{3, 4}. According to these scholars, only years of immersion into a musical culture could enable us to truly understand its uniqueness. However, embracing cultural uniqueness leaves little room for cross-cultural comparability. If instead we consider cross-cultural comparisons, illuminating patterns like the so-called *statistical universals* can emerge. These statistical universals are musical features that appear above chance across many or all musical cultures in the world^{5, 6}. At the other extreme, experimental psychologists and cognitive neuroscientists have traditionally focused on *musicality* by probing into the biocognitive substrates that enable music processing^{7, 8}. This neuropsychological approach assumes that these biocognitive substrates are common to all humans. As a result, data from a specific population and culture is assumed to be generalizable to mankind at large. While exceptions exist⁹, the field of music research has largely been polarized into these two apparently segregated dimensions. These separate approaches to music research can be unified and revelatory, especially when asking evolutionary questions^{10, 11}.

How music changes over time seems to be orthogonal to the evolution of the biocognitive apparatus underlying music-making. Recent experiments, however, have shown that music and musicality are intimately connected¹²⁻¹⁶. Focusing on rhythm, individual participants¹⁴ were asked to imitate snare-drum sequences on an electronic drum set to the best of their abilities. All participants were non-musicians, and were not told where the patterns came from. A ‘first generation’ of participants was given computer-generated drum patterns, featuring random beat intensity and duration. Successive ‘generations’, however, were asked to imitate the output of the previous participant in the experiment (Figure 1A, top ‘chain’ of participants). Unbeknownst to each participant, patterns included all the errors and imperfections introduced in the previous generation. In this way, the process of cultural transmission of rhythmic patterns was recreated

1
2
3 in the lab^{12, 17, 18}. Over time, these sequences became more structured and easier to imitate. In
4 addition, rhythmic sequences converged towards all rhythmic universal features which
5 characterize music around the world^{6, 19}. This experiment, which aligns with similar findings in
6 non-musical domains²⁰⁻²⁵, suggests a link among cognition, biology, and culture in human
7 music^{26, 27}. It shows that features present in almost all musical traditions around the world
8 emerge through basic bio-cognitive biases, and are thereby amplified by the process of cultural
9 transmission. This experiment also raised a number of additional questions. In particular, are
10 the bio-cognitive biases - responsible for emergence of structure and universals - specific to: (1)
11 all humans, (2) adults, (3) Westerners, (4) a cognitive domain or modality, and (5) particular
12 individuals?
13
14
15
16
17
18
19
20

21 This paper begins to tackle the question: how do individual biases affect the process of cultural
22 evolution of music²⁸? Two alternative hypotheses - originally advanced for language - can be
23 readily adapted^{27, 29}. One 'nativist' hypothesis sees human cognition as constrained to produce
24 only certain forms of music, and musical structure as 'designed' by individual minds and
25 transmitted down generations. Another 'interactive' hypothesis is that the structure in music
26 results from long-term distributed effects of multiple, individually varying minds that together
27 create a kind of structural compromise that is pleasing to and learnable by all^{17, 18}. These
28 contrasting hypotheses have an exact parallel in the human language and animal
29 communication literatures^{30, 31 32, 33}. In the nativist scenario, if all individuals share strongly
30 constraining biases, music would emerge as a result of strengthening few, human-widespread
31 tendencies towards musicality (akin to a vote by consensus). In the alternative interactive
32 scenario, if all individuals have weak and potentially variable biases, music would emerge as a
33 result of the interplay among individual-specific tendencies towards musicality, amplifying or
34 averaging each other out (similarly to mixing paint of different colours). To summarize, we ask:
35 do regularities in musical structure result (a) from specific individuals who impose universal
36 structural regularities, or (b) from distributed, weakly-biased processing³⁴?
37
38
39
40
41
42
43
44
45
46
47
48
49

50 We address this question using experimental manipulations and Bayesian modeling techniques.
51 First, we replicate a previous experiment with one variation¹⁴. In the original experiment, a set of
52 drumming patterns was transmitted across generations of participants in a between-participants
53 design (top row of Figure 1A). In contrast, the current experimental design features a within-
54
55
56
57
58
59
60

participant structure (bottom two rows of Figure 1A)^{13, 30, 34, 35}. Each participant takes part in multiple rounds, instead of one round per participant as in the previous experiment. Second, we analyze and compare the data from the two experiments using standard inferential statistics. Third, we introduce a probabilistic model for latent structures underpinning rhythmic sequences, alongside a psychologically plausible algorithm for inferring these structures. This allows us to obtain approximate structural descriptions of rhythmic patterns across conditions and generations, and to explore how these structures are used and re-used. Whereas previous models have focused on inferring cognitive biases for integer ratio rhythmic categories from experimental data¹³, our model focuses on approximating the process through which individuals combine rhythmic categories into predictable motif-like sequences.

A general prediction is that the new within-participants “chains” will likely produce data *qualitatively* comparable to the original between-participant design. In particular, patterns should increase in structure, and become easier to learn. If this holds, self-learning will be shown as an effective method for uncovering musical biases in participants. *Quantitatively*, however, the two experimental designs might show different behaviours. In particular, if participants have individual-specific biases, they will likely impose an idiosyncratic structure every generation. This, in turn, will make structure emerge faster in the within-participant chains, and slower in between-participants chains (i.e. some innovations will cancel out). If this is the case, we should observe a significant difference between the variables measuring the evolution of structure in the two experiments. If, instead, participants have homogeneous biases towards rhythmic structures, the evolution of chains will be unaffected by experimentally substituting many participants for one repeated participant. In this case, the two experimental designs will produce similar data, [hence no measurable difference of key variables between conditions](#).

Experimental Methods

Data from six experimental chains (30 participants) from a previous experiment were reanalyzed (details in ¹⁴). In brief, six different sets of 32 sequences of 12 random beats were given to ‘first generation’ participants to imitate. First generation output became second generation input, and so on (see Figure 1). In addition, new data was collected from 12 experimental chains (12 participants) in conditions comparable to the previous experiment¹⁴, with two key differences (for

details see Supplement). First, each of the six different sets of 32 sequences of 12 random beats used in the first experiment was given to two first-generation participants in this new experiment (as opposed to one first-generation participant in the previous experiment, see Figure 1). Second, the new experiment featured within-individual transmission, so that the same participant listened and imitated their own drum patterns over 5 experimental ‘generations’.^b

Frequentist statistics: Results and discussion

The metrics tracking structure and imitation error behave similarly over generations (Figure 1B, 1C), confirming our qualitative prediction. ANOVAs tested whether the ‘generation’ and ‘transmission type’ (i.e., between or within-individual) could account for a possible increase in structure^{14, 36, 37} (Figure 1B) and decrease in imitation error^{14, 38} (Figure 1C). Stepwise model selection suggested that both generation and transmission type should be entered in the ANOVA as predictors of structure (minimizing Akaike Information Criterion). Both variables were significant predictors of structure (transmission type: $F=7.4$, $p<.01$; generation: $F=14.5$, $p<.001$). Another stepwise model selection suggested that only generation should be entered in the ANOVA as predictor of imitation error. Generation was a significant predictor of imitation error ($F=14.8$, $p<.001$).

We used further ANOVAs to test that these differences were not due to superficial features. We calculated the length of each drumming pattern (in the ‘r space’, see Figure 2), and computed mean and variance length for each set of 32 patterns. Stepwise model selection suggested that neither generation nor transmission type should be entered in the ANOVA as mean pattern length, and only generation should be entered in the ANOVA as predictor of variance in pattern length. Generation was a significant predictor of variance in pattern length ($F=4.0$, $p<.05$). In other words, while participants vary the number of beats produced within an experimental session across generations, this does not appear affected by the transmission type. Crucially, transmission type affects neither mean nor variance in pattern length, suggesting that simple differences in pattern length cannot account alone for structural variability across experimental groups.

Together, these results suggest that both structure and learnability change over transmission steps. Although transmission type does not affect learnability, it does affect the amount of structure: a within-participant design results in higher levels of structure. This provides preliminary support for the ‘idiosyncrasy of biases’ hypothesis over the ‘homogeneity of biases’ hypothesis. These inferential statistics suggest that repeated idiosyncrasies enhance the emergence of structure (within-participant transmission type). Conversely, new minds introduce more variance, slowing down the emergence of structure (between-participants transmission type). Inferential statistics, however, are unfit to unveil the structures participants infer. Below we outline a mathematical model approximating what participants perceive and learn.

Bayesian model: Methods

Data from both experiments were further analyzed using a computational model (see Supplement). Our model formalises the idea that participants may decompose individual drum patterns into sequences of motifs that can be reused across patterns. Given a drum pattern, the model attempts to infer boundaries between latent motifs, and to categorise these motifs into coherent groups based on prototypes. Our approach takes inspiration from two related fields. First, since the inferential task is essentially one of joint segmentation and clustering, we can adapt techniques from the machine learning and speech technology literature³⁹ to specify a probabilistic model of underlying latent variables (e.g. motifs and their boundaries). Second, the literature on statistical learning provides a psychologically-plausible algorithm that makes guesses about these unobserved variables⁴⁰.

Our model formalises two levels of structure within a pattern, by grouping adjacent interval ratios⁴¹ into a small number of Gaussian categories via Bayesian inference^a, and by using sequences of these inferred categories to construct an inventory of motifs that can be reused. Our posterior approximation algorithm learns this structure ‘on the fly’ in a probabilistic, sequentially-dependent, psychologically-plausible fashion: as new elements of a drumming sequence are perceived, the backwards transition probabilities (BTPs) between interval ratio categories are estimated³⁷, and used to hypothesise boundaries between motifs. The model attempts to assign any hypothesised motif to an existing category of motifs (Figure 2, top), and creates a new motif category whenever this fails (Figure 2, bottom).

1
2
3
4
5
6 A candidate subsequence has to go through two criteria to be considered a motif: (1) how often
7 has the model seen this particular subsequence of categories? (2) how often have the
8 preceding and current element been found together (BTPs)? If the same sequence has been
9 seen before, this provides evidence for the current subsequence to be another occurrence of
10 this motif type. If the first element of a subsequence and the previous element rarely co-occur,
11 this provides evidence that the two elements belong to different motifs, implying a boundary.
12 The model includes parameters that influence, for example, how willing the learner is to invent
13 new motif categories. These parameters and all other details of the model are described in
14 detail in the supplement to this article. In the analyses we present here, we set these
15 parameters such that the model is weakly biased to prefer re-using existing motifs, but able to
16 invent new categories whenever the data dictate. Because the model includes a free parameter
17 that determines this balance of re-use and invention, this assumption could be straightforwardly
18 revisited in future analyses. Crucially, we fix parameters to be identical across analyses of both
19 experimental datasets, and examine how the structures inferred by the model vary by
20 experimental condition (rather than examine the specific structures inferred, which can be
21 sensitive to model parameterization).
22
23
24
25
26
27
28
29
30
31
32
33

34 **Bayesian model: Results and discussion**

35
36 We ran the model through both experimental datasets 10 times, which (insofar as our model is a
37 psychologically-plausible theory of participants' behaviour) provides an approximation to the
38 representation of patterns induced by participants. We examined two principal measures of
39 structure: 1) the number of unique motifs discovered by the model at each generation of each
40 chain, and 2) the number of patterns in which each attested motif was discovered at each
41 generation of each chain. Together, these measures quantify the evolution of structural
42 regularity across a set of rhythms, over generations, as a function of the data participants saw
43 and produced.
44
45
46
47
48
49
50
51

52 In both experimental conditions, the number of unique motifs attested within a generation
53 decreased over generations (Figure 3A, top row). In line with our inferential statistics, this
54 suggests an increased degree of re-use of prototypical building blocks over generations, and
55
56
57
58
59
60

that these building blocks are discoverable by a simple algorithm making local, sequential decisions. The regression slope plotted in Figure 3A shows that re-use happens faster in the within-subjects chains, and that the final generation of within-subjects chains re-uses motifs to a slightly greater extent than the final generation of between-subjects chains.

Our analyses also suggest (Figure 3A, second row) that in within-subjects chains, each attested motif tends to be present in an increasing number of patterns over generations, suggesting participants are entrenching the motifs they have invented. This pattern is also visible in the between-subjects chains, but to a lesser extent. While these chains do evidence increasing re-use of motifs, they do not appear to evidence the same degree of entrenchment on a small set of widely re-used motifs as is suggested in the within-subjects results. We interpret this as a sign that (1) both experimental conditions lead to an increase in structure, but (2) the within-subjects condition allows the idiosyncrasies of individual minds to repeatedly bias the distribution of structures in a chain-specific way that is less probable when new learners are forced to re-interpret the structures invented by previous individuals.

Visualizing integer ratios

We computed the distribution of interval ratios^{13, 14} at each generation (Figure 3B). Both experimental conditions evidence a sharp transition from unstructured initial distributions (generation 0, top row) to highly structured categories of intervals. Interestingly, while the tri-modal distribution with peaks near interval ratios is clear in the between-subjects data (as previously reported¹⁴), the within-subjects chains appear to converge on an approximately two-way category distribution with peaks at 1:2 and 1:1 ratios^{13, 42}. Figure 4 illustrates these differences in more detail, providing an overview of the evolution of interval ratio distributions across generations. The figure shows Gaussian Kernel density estimates of the distribution of interval ratios that participants observed (blue lines) and then produced (black lines) at each generation of each chain (in both conditions).

Focusing on the between-participants chains (Figure 4A), all five chains independently converge on an approximately three-way category distinction with one category of roughly equal-length

intervals, and one category each for long-short intervals and short-long intervals. In contrast, the within-participant chains (Figure 4B) do not converge so emphatically towards this musical universal^{6, 13, 14}. Instead, within-participant chains demonstrate idiosyncratic (but nonetheless structured) final distributions. Only two of the within-participant chains (chains 3 and 6) appear to approximate the 3-way category distinction found in all between-participant chains, and even here those distinction are less clear. The other four within-participant chains converge on different solutions: a single primary category of equal length intervals strongly peaked at zero (chain 1); a four-way distinction with two below-zero categories (chain 2); a two-category distinction (chain 4) with roughly equal-length intervals and long-short intervals; and an approximately five-way distinction covering the range (chain 5). It is also notable that the majority of chains in both conditions include a well-defined final category at mean zero – roughly adjacent equal-length intervals.

Also note the dynamics of distribution change over generations. In the between-participant chains, change is gradual and constant until the final generations: the distribution of produced interval ratios (black line) marginally deviates from the observed distribution (the black line). In contrast, in the within-participant chains, we see an initial generation of deviance between observed and produced distributions as in the between-participant chains, but the following generations generally reproduce the observed distribution accurately (black lines trace blue lines). Along with the modelling results above, we take this to imply that within-participant chains do result in structured rhythms, but that this structure is largely idiosyncratic, arguing strongly against the nativist explanation for these universals.

General discussion and conclusions

In this paper, we investigate the role of individual cognitive biases in the creation of rhythmic patterns. We present two rhythm-imitation experiments (simulated cultural transmission) in non-musicians, analyzed using two complementary techniques. We show that similar regularities emerge when participants are asked to imitate their own or other participants' patterns. However, different designs affect the amount of regularities emerging: When participants imitate their own previous productions, convergence is faster but results in less pronounced universal rhythmic features. We suggest this is due the presence of *weak idiosyncratic biases*. When transmission occurs between participants, idiosyncratic biases partially cancel each other out. Instead, when transmission occurs within a participant, biases reinforce each other.^c

Previous cultural transmission experiments of musical rhythms show that^{12, 14-16}: (1) initially random sequences become structured; (2) the resulting structures reflect universally observed patterns in music, such as small integer ratios of durations between note onsets. Our experiment probed whether repeated computation by the same individual leads – as in the original experiment¹⁴ – to increased structure and increased proximity to rhythmic universals. Our findings are surprising because (1) structure indeed emerges, and even *faster* than in the between-participants chains, but (2) the resulting structure appears to be a poor approximation to the integer ratio universal^{6, 13, 14}. We take this to be supportive of an account in which rhythmic structure results from a balance between individuals imprinting their biases, and interactive transmission among listeners, learners, and performers. Our results support the interactive hypothesis over the nativist hypothesis²¹.

Several limitations of this study could be addressed by future work. First, we adapted and designed the model to have cognitive plausibility and match experimental conditions as closely as possible. However, our Bayesian model still neglects several findings from rhythm perception and production^{43, 44}, and should be understood as a first approximation. Future research should refine the model towards greater psychological and neural plausibility. Likewise, while we focused on psychological plausibility by implementing a sequential posterior approximation algorithm as our model of the learner, future analyses might instead focus on inferring the best possible structural descriptions, by implementing more computationally intensive posterior sampling procedures. Second, the intensity information of each beat is unused in our models. Velocity and accents are an integral part of rhythm, so future extensions of this work should go beyond purely durational information. Third, several parameters in the Bayesian model undoubtedly influence our results. Future modelling efforts might aim to find empirical motivation for parameters accounting for human perception and induction of categories of intervals^{43, 44}, or derive simpler models with fewer parameters. Ours is a first attempt, and we will share our data and scripts with interested researchers who would like to perform modifications. Fourth, given a restricted set of assumptions, we have predicted the number and distribution of rhythmic motifs inferred by participants. Ideally, whether participants actually acquire similar motif-like substructures will need to be tested experimentally by asking participants to classify motifs, and check how closely these decisions align with the model's predictions.^d

1
2
3
4
5
6 This report makes a contribution to a number of disciplines. Over the past century, a deep divide
7 has separated cultural anthropology and music psychology^{1, 11}. Our experiments aim to bridge
8 this divide with a design accounting for both the cultural medium and human bio-psychological
9 features. Likewise, our behavioural experiments could be combined with neuroimaging or
10 electrophysiology techniques⁴⁵ to tap into the neural basis of human biases for musical rhythm.
11 Within the interdisciplinary field of cultural evolution, we show how within-participant
12 transmission speeds up the process of convergence. Quantitative models abound in music
13 information retrieval, but are still scarce in music cognition. Here, we adapted some recent
14 computational techniques to the interpretation of human data. In brief, we hope that our paper
15 will spur a tighter integration of modelling and empirical research in the study of the psychology,
16 neuroscience and evolution of music.
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Acknowledgments

AR and BT are grateful to Marieke Woensdregt and Carmen Saldaña Gascón for hospitality and support during the development of this research. Andrea Ravignani has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 665501 with the research Foundation Flanders (FWO) (Pegasus² Marie Curie fellowship 12N5517N), and a visiting fellowship in Language Evolution from the Max Planck Society.

Authors contributions: AR and SK conceived the research, AR, TG, TD and SK designed the experiments, TG and TD performed the experiments, all authors analyzed the data, AR, BT and SK conceived the model, BT implemented the model, AR and BT wrote the manuscript, all authors edited and approved the manuscript for publication.

Conflicts of interest

The authors declare no conflicts of interest.

Figure legends:

Figure 1. Experimental design and summary statistics tracking the evolution of patterns in the first 6 generations of both experiments. (A) The experimental design follows a transmission chain paradigm: the output of one ‘experimental generation’ constitutes the input of the next generation. Generation 0 consists of computer-generated drumming patterns, where drum hits have random velocity and inter-onset intervals (IOIs). Each generation 0 pattern is individually heard and imitated by a Generation 1 participant. The resulting imitated patterns can be given to the same participant to imitate once again (bottom two rows of A, within-individual design), or to a different participant (top row of A, between-individuals design). The procedure is repeated over generations (left to right), and additional chains (not shown). (B) Increase in structural complexity quantified over generations using a modified measure of entropy. (C) Decrease in imitation error between adjacent generations, quantified using a modified Levenshtein distance. (D) Pattern length across generations and chains. (E) Variance in pattern length across generations and chains. (B-E) Shaded areas depict bootstrapped 95% confidence intervals, lines connect the mean value for each generation and experiment type across chains.

Figure 2. Latent variables Bayesian model: Sketch of how the algorithm processes two drumming patterns (workflow proceeds from top to bottom, and from left to right). A drumming pattern (first pattern, topmost row) can be conceptualized as a series of duration marked by drum events (lines broken by circles). The (absolute) durations between drum events are represented as a vector of IOIs. By taking the ratios between adjacent IOIs, one obtains a vector of relative durations $r=(r_1, \dots, r_{n-2})$. All IOI sequences with the same r vector have the same rhythmic pattern up to a tempo multiplicative constant. The algorithm generalizes first over r_i categories (e.g. in the first pattern .5 and .5 belong to the same category) and then assigns every hypothesized motif either to its prototypical category or to a new category. When the participant hears a new pattern (second pattern) with more variability, ratios such as .54 and .52 might be assigned to the same r_i category. Likewise, the algorithm randomly attempts to be 'greedier', probing the existence of motifs of length 3 or above, hence finding that sequences like (.54, 1, 2) and (.52, 1, 2.1) belong to the same category.

Figure 3: (A) The number of motifs inferred by the model (averaged over 10 independent simulations) at each generation (top row) in the within-subjects (right) and between-subjects (left) chains, and the number of independent patterns in which each motif evidenced at least once in a generation was identified in that generation (bottom row). Points show these quantities for individual motifs (simulated ten times); lines show regression slopes. **(B)** Normalised histograms for the distribution of interval ratios at each generation in the between-subjects (right) and within-subjects (left) chains. Lines show kernel density estimates of these distributions.

Figure 4. The evolution of interval-ratio distributions over generations, for all chains, in the between-participants (A) and within-participants experiments (B). Lines show Gaussian kernel density estimates inferred from (the logarithms of) the raw interval ratios participants observed (blue lines) and produced (black lines). We plot the distribution of the *logarithm* of interval ratios since we found this to be the clearest illustration of categories. An interval ratio of 1 denotes adjacent equal-length intervals, and the logarithm of this ratio is 0. Values below zero in log space indicate that the second interval is shorter than the first (long-short interval); values above zero indicate that the second interval is longer (short-long interval).

Footnotes and Endnotes: (Use lower-case italic letters in superscript.)

^aEach of the 32 imitated pattern can be described as a time series of inter-onset intervals, i.e. the time between adjacent drum hits $IOI_1, IOI_2, \dots, IOI_n$. As in the original study, to account for possible tempo drift within and across patterns, we use the ratio between adjacent beats^{41, 46}, i.e. $r_i = IOI_{i+1}/IOI_i$. Hence each pattern of n hits can be represented as a time series r_1, \dots, r_{n-2} . The computational model proceeds by first clustering data points into rhythmic categories. We assume that participant do not perceive and represent the absolute magnitude of the r_i data points, but potentially reduce the variation among data points by assigning each to a rhythmic category⁴³.

^bIn both experiments, the participant was unaware that she would imitate her own, or someone else's, previous pattern.

^cOur conclusions partly contrast with computer simulations done for language, where a few outlier agents distort the signal transmitted in an otherwise heterogeneous population³⁴.

^dWhile we would expect less difference between conditions if priors were strong and homogenous, the design of the two experiments differs along one additional dimension. In the within-participant experiment, participants can potentially carry memory over from previous generations. If that were the case, one would predict a slower evolution of structure in the within-participant condition, as the effect of the prior would be relatively weakened from the drag of increasing pile of data. The fact that this does not occur may suggest that - at least in this design - participant memory is not a crucial factor. In fact, subjects in the within-participant condition were not told they were listening to their own data. That, plus the sheer number and length of sequences, and the participants' lack of musical training should explain why memory is not as strong a factor. Future research will be faced with disentangling bias' homogeneity from memory effects. Assigning participants to experimental groups based on their electroencephalographic signature might be a fruitful solution⁴⁵.

References (no more than 100. cited in the text in numerical order as numbers)

1. Ravignani, A., B. Thompson & P. Filippi. in review. The evolution of musicality: What is to learn from language evolution research? *Frontiers in Neuroscience*.
2. Honing, H., C. ten Cate, I. Peretz, *et al.* 2015. Without it no music: cognition, biology and evolution of musicality. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*. **370**: 20140088.
3. Arom, S. 1991. *African polyphony and polyrhythm: musical structure and methodology*. Cambridge university press Cambridge.

4. Nettl, B. 2000. "An ethnomusicologist contemplates universals in musical sound and musical culture". In *The Origins of Music*. N.L. Wallin, B. Merker & S. Brown, Eds.: 463-472. Cambridge, Mass.: The MIT Press.
5. Brown, S. & J. Jordania. 2013. Universals in the world's musics. *Psychology of Music*. **41**: 229-248.
6. Savage, P.E., S. Brown, E. Sakai, *et al.* 2015. Statistical universals reveal the structures and functions of human music. *Proceedings of the National Academy of Sciences*. **112**: 8987-8992.
7. Peretz, I., D. Vuvan, M.-â. Lagrois, *et al.* 2015. Neural overlap in processing music and speech. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*. **370**: 20140090.
8. Sameiro-Barbosa, C.M. & E. Geiser. 2016. Sensory Entrainment Mechanisms in Auditory Perception: Neural Synchronization Cortico-Striatal Activation. *Frontiers in neuroscience*. **10**.
9. Fritz, T., S. Jentschke, N. Gosselin, *et al.* 2009. Universal recognition of three basic emotions in music. *Current Biology*. **19**: 573-576.
10. Honing, H. & A. Ploeger. 2012. Cognition and the evolution of music: Pitfalls and prospects. *Topics in cognitive science*. **4**: 513-524.
11. Lumaca, M., A. Ravignani & G. Baggio. in review. Music evolution in the laboratory: Cultural transmission meets neurophysiology. *Frontiers in Neuroscience*.
12. Fitch, W.T. 2017. Cultural evolution: Lab-cultured musical universals. *Nature Human Behaviour*. **1**: 1-2.
13. Jacoby, N. & J.H. McDermott. 2017. Integer Ratio Priors on Musical Rhythm Revealed Cross-culturally by Iterated Reproduction. *Current Biology*.
14. Ravignani, A., T. Delgado & S. Kirby. 2016. Musical evolution in the lab exhibits rhythmic universals. *Nature Human Behaviour*. **1**: 0007.
15. Ravignani, A. & T. Verhoef. in press. Which melodic universals emerge from repeated signaling games? *Artificial Life*.
16. Lumaca, M. & G. Baggio. 2017. Cultural transmission and evolution of melodic structures in multi-generational signaling games. *Artificial Life*.
17. Le Bomin, S., G. Lecointre & E. Heyer. 2016. The Evolution of Musical Diversity: The Key Role of Vertical Transmission. *PLoS one*. **11**: e0151570.
18. Trehub, S.E. 2015. Cross-cultural convergence of musical features. *Proceedings of the National Academy of Sciences*. **112**: 8809-8810.

19. Rzeszutek, T., P.E. Savage & S. Brown. 2012. The structure of cross-cultural musical diversity. *Proceedings of the Royal Society of London B: Biological Sciences*. **279**: 1606-1612.

20. Verhoef, T., S. Kirby & B. de Boer. 2014. Emergence of combinatorial structure and economy through iterated learning with continuous acoustic signals. *Journal of Phonetics*. **43**: 57-68.

21. Thompson, B., S. Kirby & K. Smith. 2016. Culture shapes the evolution of cognition. *Proceedings of the National Academy of Sciences*. **113**: 4530-4535.

22. Miranda, E.R., S. Kirby & P. Todd. 2003. On computational models of the evolution of music: From the origins of musical taste to the emergence of grammars. *Contemporary Music Review*. **22**: 91-111.

23. Kirby, S., T. Griffiths & K. Smith. 2014. Iterated learning and the evolution of language. *Current Opinion in Neurobiology*. **28**: 108-114.

24. Kirby, S., M. Dowman & T.L. Griffiths. 2007. Innateness and culture in the evolution of language. *Proceedings of the National Academy of Sciences*. **104**: 5241-5245.

25. Kirby, S., H. Cornish & K. Smith. 2008. Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences*. **105**: 10681-10686.

26. Ravignani, A., H. Honing & S.A. Kotz. 2017. The evolution of rhythm cognition: Timing in music and speech. *Frontiers in human neuroscience*. **11**.

27. Kirby, S. 2017. Culture and biology in the origins of linguistic structure. *Psychonomic bulletin & review*. **24**: 118-137.

28. Gingras, B., H. Honing, I. Peretz, *et al.* 2015. Defining the biological bases of individual differences in musicality. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*. **370**: 20140092.

29. Berwick, R.C. & N. Chomsky. 2017. Why only us: Recent questions and answers. *Journal of Neurolinguistics*. **43**: 166-177.

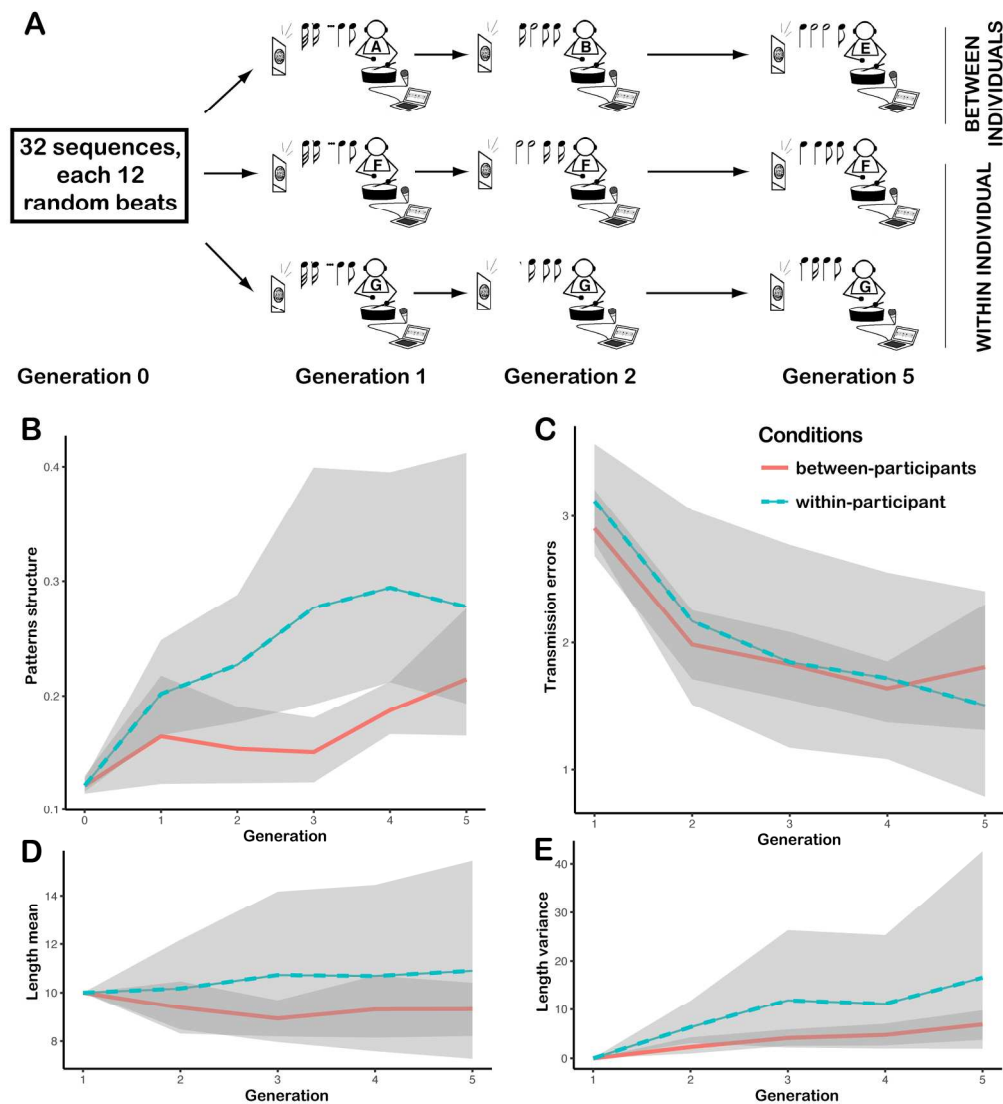
30. Fehér, O., I. Ljubičić, K. Suzuki, *et al.* 2017. Statistical learning in songbirds: from self-tutoring to song culture. *Phil. Trans. R. Soc. B*. **372**: 20160053.

31. Fehér, O., H. Wang, S. Saar, *et al.* 2009. De novo establishment of wild-type song culture in the zebra finch. *Nature*. **459**: 564-568.

32. James, L.S. & J.T. Sakata. 2017. Learning Biases Underlie “Universals” in Avian Vocal Sequencing. *Current Biology*.

33. Tchernichovski, O. & D. Lipkind. 2017. Animal Communication: Origins of Sequential Structure in Birdsong. *Current Biology*. **27**: R1268-R1269.

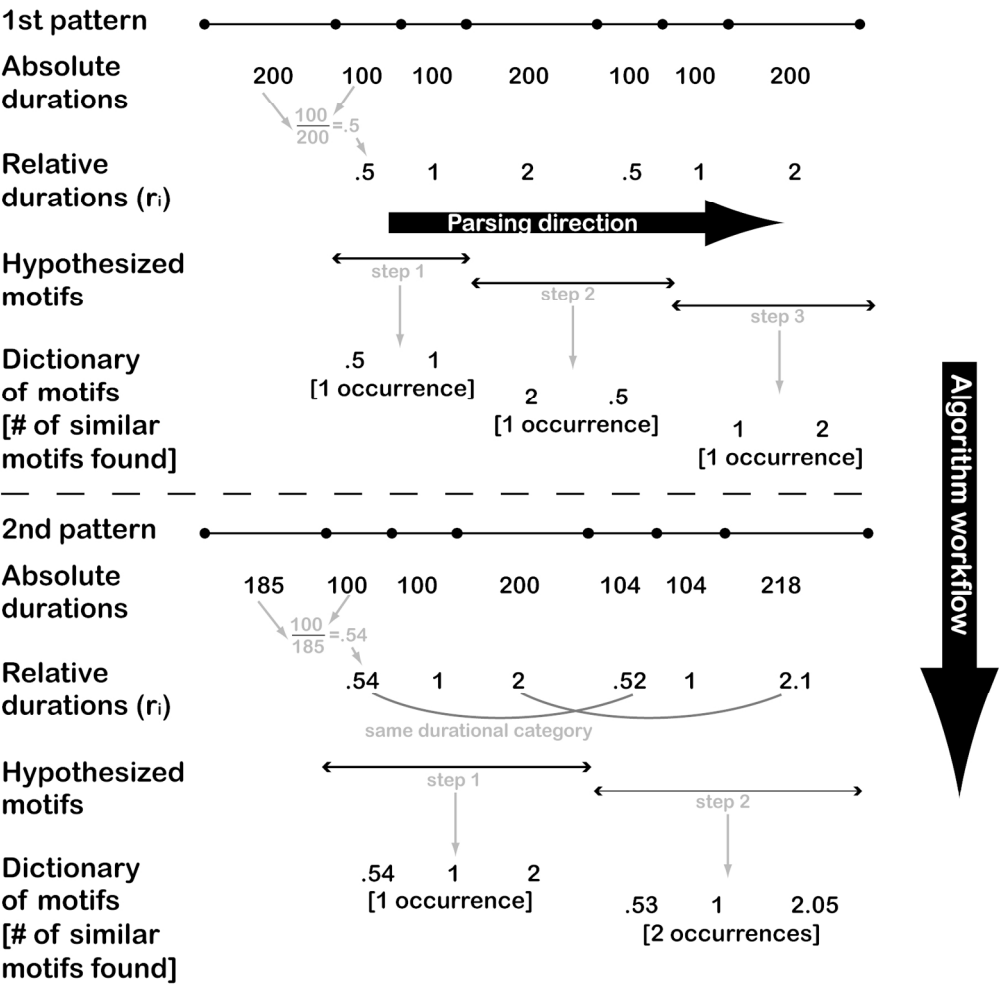
34. Navarro, D.J., A. Perfors, A. Kary, *et al.* "When extremists win: On the behavior of iterated learning chains when priors are heterogeneous". In *CogSci 2017*.
35. Griffiths, T.L. & M.L. Kalish. 2007. Language evolution by iterated learning with Bayesian agents. *Cognitive science*. **31**: 441-480.
36. Jamieson, R.K. & D. Mewhort. 2009. Applying an exemplar model to the artificial-grammar task: Inferring grammaticality from similarity. *The Quarterly Journal of Experimental Psychology*. **62**: 550-575.
37. Ravignani, A. & G. Madison. 2017. The paradox of isochrony in the evolution of human rhythm. *Frontiers in psychology*.
38. Levenshtein, V.I. 1966. "Binary codes capable of correcting deletions, insertions, and reversals". In *Soviet physics doklady*, Vol. 10: 707-710.
39. Lee, C.-y. & J. Glass. 2012. "A nonparametric Bayesian approach to acoustic model discovery". In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Long Papers-Volume 1*: 40-49. Association for Computational Linguistics.
40. McCauley, S.M. & M.H. Christiansen. 2011. "Learning simple statistics for language comprehension and production: The CAPPUCINO model". In *Proceedings of the Cognitive Science Society*, Vol. 33.
41. Sakai, K., O. Hikosaka, S. Miyauchi, *et al.* 1999. Neural representation of a rhythm depends on its interval ratio. *Journal of Neuroscience*. **19**: 10074-10081.
42. Motz, B.A., M.A. Erickson & W.P. Hetrick. 2013. To the beat of your own drum: Cortical regularization of non-integer ratio rhythms toward metrical patterns. *Brain and cognition*. **81**: 329-336.
43. Desain, P. & H. Honing. 2002. The formation of rhythmic categories and metric priming. *Perception*. **32**: 341-365.
44. Honing, H. & D. Deutsch. 2013. Structure and interpretation of rhythm in music. *The psychology of music*. 369-404.
45. Lumaca, M. & G. Baggio. 2016. Brain potentials predict learning, transmission and modification of an artificial symbolic system. *Social cognitive and affective neuroscience*. **11**: 1970-1979.
46. Grondin, S. 2010. Timing and time perception: a review of recent behavioral and neuroscience findings and theoretical directions. *Attention, Perception, & Psychophysics*. **72**: 561-582.



Experimental design and summary statistics tracking the evolution of patterns in the first 6 generations of both experiments. (A) The experimental design follows a transmission chain paradigm: the output of one 'experimental generation' constitutes the input of the next generation. Generation 0 consists of computer-generated drumming patterns, where drum hits have random velocity and inter-onset intervals (IOIs). Each generation 0 pattern is individually heard and imitated by a Generation 1 participant. The resulting imitated patterns can be given to the same participant to imitate once again (bottom two rows of A, within-individual design), or to a different participant (top row of A, between-individuals design). The procedure is repeated over generations (left to right), and additional chains (not shown). (B) Increase in structural complexity quantified over generations using a modified measure of entropy. (C) Decrease in imitation error between adjacent generations, quantified using a modified Levenshtein distance. (D) Pattern length across generations and chains. (E) Variance in pattern length across generations and chains. (B-E) Shaded areas depict bootstrapped 95% confidence intervals, lines connect the mean value for each generation and experiment type across chains.

182x199mm (300 x 300 DPI)

unedited manuscript



Latent variables Bayesian model: Sketch of how the algorithm processes two drumming patterns (workflow proceeds from top to bottom, and from left to right). A drumming pattern (first pattern, topmost row) can be conceptualized as a series of duration marked by drum events (lines broken by circles). The (absolute) durations between drum events are represented as a vector of IOIs. By taking the ratios between adjacent IOIs, one obtains a vector of relative durations $r=(r_1,...,r_{n-2})$. All IOI sequences with the same r vector have the same rhythmic pattern up to a tempo multiplicative constant. The algorithm generalizes first over r_i categories (e.g. in the first pattern .5 and .5 belong to the same category) and then assigns every hypothesized motif either to its prototypical category or to a new category. When the participant hears a new pattern (second pattern) with more variability, ratios such as .54 and .52 might be assigned to the same r_i category. Likewise, the algorithm randomly attempts to be 'greedier', probing the existence of motifs of length 3 or above, hence finding that sequences like (.54, 1, 2) and (.52, 1, 2.1) belong to the same category.

152x149mm (300 x 300 DPI)

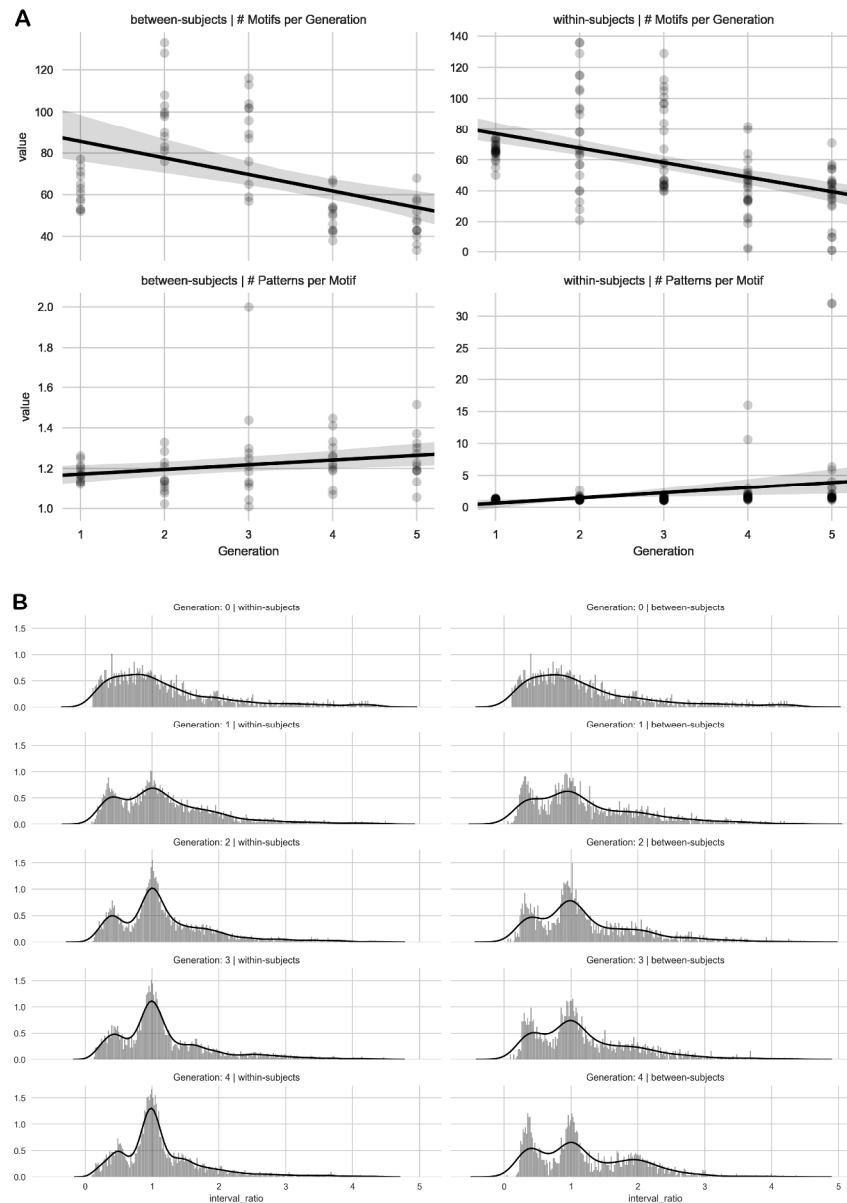


Figure 3: (A) The number of motifs inferred by the model (averaged over 10 independent simulations) at each generation (top row) in the within-subjects (right) and between-subjects (left) chains, and the number of independent patterns in which each motif evidenced at least once in a generation was identified in that generation (bottom row). Points show these quantities for individual motifs (simulated ten times); lines show regression slopes. (B) Normalised histograms for the distribution of interval ratios at each generation in the between-subjects (right) and within-subjects (left) chains. Lines show kernel density estimates of these distributions.

296x417mm (300 x 300 DPI)

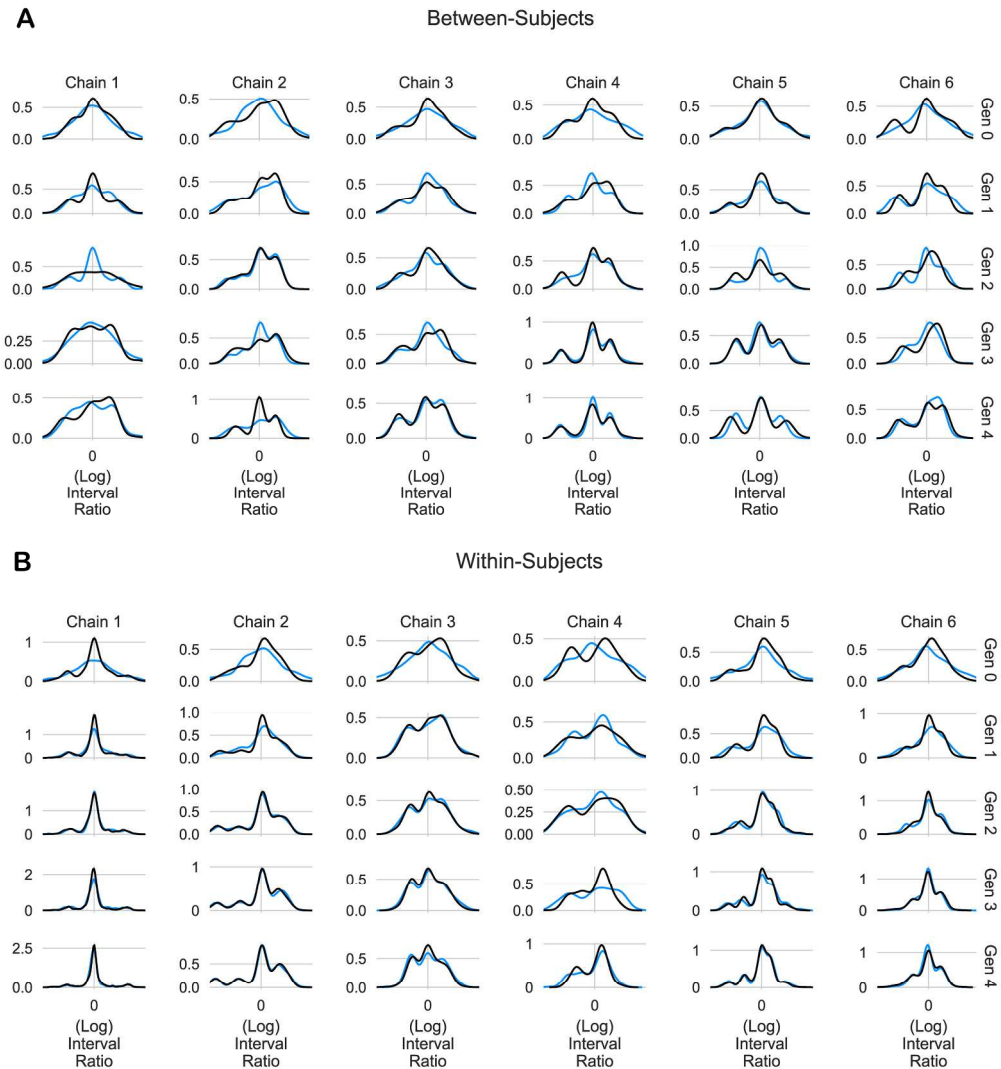


Figure 4. The evolution of interval-ratio distributions over generations, for all chains, in the between-participants (A) and within-participants experiments (B). Lines show Gaussian kernel density estimates inferred from (the logarithms of) the raw interval ratios participants observed (blue lines) and produced (black lines). We plot the distribution of the logarithm of interval ratios since we found this to be the clearest illustration of categories. An interval ratio of 1 denotes adjacent equal-length intervals, and the logarithm of this ratio is 0. Values below zero in log space indicate that the second interval is shorter than the first (long-short interval); values above zero indicate that the second interval is longer (short-long interval).

227x243mm (300 x 300 DPI)